

# UNCLASSIFIED

<b>AD NUMBER</b>
ADB011564
<b>NEW LIMITATION CHANGE</b>
<b>TO</b> Approved for public release, distribution unlimited
<b>FROM</b> Distribution limited to U.S. Gov't. agencies only; Test and Evaluation; May 76. Other requests for this document must be referred to Director, Naval Research Lab., Washington, D. C. 20375.
<b>AUTHORITY</b>
NRL ltr., Ser 96-018, 5 Jan 1996

THIS PAGE IS UNCLASSIFIED

NRL Report 7989

# Track-to-Track Association in Ocean Surveillance

HOWARD L. WIENER

*Systems Research Branch  
Space Systems Division*

May 12, 1976

AD No. \_\_\_\_\_  
DDC FILE COPY



DDC  
RECEIVED  
JUN 11 1976  
D

NAVAL RESEARCH LABORATORY  
Washington, D.C.

Distribution limited to U.S. Government Agencies only; test and evaluation; May 1976. Other requests for this document must be referred to the Director, Naval Research Laboratory, Washington, D.C. 20375.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL-7989	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TRACK-TO-TRACK ASSOCIATION IN OCEAN SURVEILLANCE	5. TYPE OF REPORT & PERIOD COVERED Final report, one phase of continuing NRL problems	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Howard L. Wiener	8. CONTRACT OR GRANT NUMBER(s)	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, D.C. 20375	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NRL Prob. BO2-26/Project 351502-75-02 NRL Prob. BO1-10 Project RRO14-02-41-6152	
11. CONTROLLING OFFICE NAME AND ADDRESS Department of the Navy Office of Naval Research Arlington, Va. 22217	12. REPORT DATE May 1976	13. NUMBER OF PAGES 31
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Distribution limited to U.S. Government Agencies only; test and evaluation; May 1976. Other requests for this document must be referred to the Director, Naval Research Laboratory, Washington, D.C. 20375.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES (16) RRO14-02, NRL-BO2-26 / (17) RRO14-02-41 /		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Surveillance, Ocean surveillance, Multisensor correlation, Multisource correlation, Track-to-track association, Statistical tests, Chi-square distribution, Noncentral chi-square distribution, Sequential tests, Mathematical algorithms, Computer programs, Requirements analysis		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The ability to resolve track-to-track association problems is one of the most important requirements in the development of automated information processors for ocean surveillance systems comprising multiple data sources. The individual tracks carried by a processor can be represented as samples from multivariate normal distributions, and the most basic track association problem—determining whether two tracks relate to the same target—is solved using a test based on the chi-square distribution. When a specified track is to be compared with two other candidates for association, a test of the chi-square distribution against a noncentral chi-square alternative is used; this		

DD FORM 1 JAN 73 1473

EDITION OF 1 NOV 65 IS OBSOLETE  
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

251950

OVER

YB

20. Abstract (Continued)

procedure can be adapted for use as a sequential test in continuing surveillance operations. The procedures developed for track association can be implemented within man-machine systems; for many cases they will make relatively few demands for interaction by surveillance analysts.

## CONTENTS

INTRODUCTION .....	1
RESULTS .....	2
OPERATIONAL SITUATION .....	2
BASIC MODEL .....	3
NULL HYPOTHESIS $H_0$ .....	5
ALTERNATE HYPOTHESIS $H_1$ .....	7
Distribution of $R^2$ Under $H_1$ ; Power of the Test .....	9
Sensitivity to Estimates of $\lambda$ .....	10
Computer Implementation .....	12
FURTHER ASPECTS OF THE PROBLEM .....	13
Nonsimultaneous Observations .....	13
Sequential Tests .....	14
Specifying the Allowable Error Probabilities .....	17
SUMMARY REMARKS .....	18
ACKNOWLEDGMENTS .....	18
REFERENCES .....	18
APPENDIX A—A Track-to-Track Association Program .....	20
APPENDIX B—Program Listing .....	23

Accession for	
NTIS	White Section <input type="checkbox"/>
DDC	Buff Section <input checked="" type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION.....	
BY.....	
DISTRIBUTION/AVAILABILITY CODES	
Dist.	AVAIL. and/or SPECIAL
B	

DDC  
**RECEIVED**  
 JUN 11 1976  
**RECEIVED**  
 D

## TRACK-TO-TRACK ASSOCIATION IN OCEAN SURVEILLANCE

### INTRODUCTION

The track-to-track association problem in ocean surveillance is discussed in this report. A family of procedures is described for deciding if two observed tracks are from the same target, and these procedures are discussed within the operational context of a large-scale automated surveillance-data processing system.

The Navy's roles in protecting and maintaining the sea lines of communication and in supporting other areas of national concern require a major effort in the surveillance of ocean surface traffic. This effort is required for monitoring possibly hostile combatant ships to protect the U.S. fleet and also for such activities as guaranteeing the freedom of movement of U.S.-bound oil tankers and locating and warning U.S.-flag merchant ships of possible dangers. The increasing need for effective and timely surveillance information has generated a requirement for diverse data sources and for the capability to process surveillance data more rapidly than can be done by unaided analysts. The Navy is developing systems which accept surveillance data of various types from a variety of sources, process the data at speeds adequate to meet increased report arrival rates, and store and maintain the finished product in a readily accessible form. A useful format for stored information is a *track*, defined as a time-ordered sequence of positions, with all sequence elements relating to the same target. Instances of the track-to-track association problem occur when two unidentified tracks are examined to determine whether they could have been developed on the same target or when two tracks identified as relating to the same target are examined to decide if the identification is, in both cases, correct.

This report is one of a series of studies resulting from continuing Naval Research Laboratory efforts in the development and analysis of surveillance systems and concepts. Recent NRL studies of surveillance problems include Refs. 1 through 5. For the case of two sensors, each of which has observed the position of a vessel, Ref. 5 developed a method for determining whether one target or two targets had actually been detected. The present study extends this situation to the case of target tracks. The test procedures discussed in this report are not unique to the track association problem. Reference 6, for example, notes that similar procedures are applicable in surveillance, quality control, and motion detection. The purpose of this report however is to describe the relevance of each of the test procedures' operating parameters to the track association problem and to indicate how these parameters will affect operational implementation of the procedures.

## HOWARD L. WIENER

### RESULTS

The major analytic results of this investigation involve the resolution of track-to-track association problems through a series of statistical tests. The basic problem of determining whether two observed tracks relate to the same target has been formulated as a test involving the chi-square distribution. An extended problem of determining which of two candidate tracks a given track can best be associated with has been formulated as a test of the chi-square distribution against a noncentral chi-square alternative. Finally the extended problem with the added feature that decisions can be deferred pending the receipt of additional data has been formulated as a sequential test. The test for the basic problem can be readily implemented within an automated surveillance-data processing system and can be carried out automatically without requiring inputs from surveillance analysts. The other tests will require inputs from analysts. The estimates required of the analyst are of operationally significant parameters, although fortunately in many cases the final decision of the test is not sensitive to moderate changes in the parameters.

### OPERATIONAL SITUATION

For this study, processing of surveillance information consists of maintaining and updating data files on tracked targets. The procedure is assumed to be carried out by a processing system which includes surveillance analysts, computers, and analyst aids, both manual and computer-driven. Computers perform the bulk of the routine data-processing functions, freeing the analysts for making decisions in difficult cases and for providing judgments which are beyond the capabilities of automated processes.

Track association schemes will be used to associate tracks resulting from a report-to-track association process. These tracks must contain the best available information on observed target movements. Report-to-track association processes operate on sets of reported target positions and sets of established tracks and attempt to match correctly the elements of these sets. The pairing of a reported position with an established track extends the track to a new position, either the position observed in the report or a smoothed position obtained from operations on the reported position and track parameters. When two or more tracks exhibit similar positions over an interval of time, a track association scheme is called into use.

The tracks upon which a track association scheme is to be used will contain the underlying positions and reported times, together with the covariance matrices associated with these positions. If the track comprises independently observed positions, then the covariance matrix will indicate no correlation between position coordinates for different observation times. However, if the track is made up of smoothed positions, then nonzero correlations will occur. This is because track smoothing schemes are based on estimating target positions by applying a computational procedure either to previous smoothed positions or directly to previous observed positions, so that each smoothed position contains information about previous smoothed positions; thus the covariance matrix associated with a set of smoothed positions will contain entries indicating correlation between the position estimates.

In this report the surveillance information is considered to be generated by two independently operating surveillance systems. The first system, called the *primary* system, is defined as possessing superior capabilities in a number of important system characteristics, such as high detection probability, high frequency of observation of a given area, or small position estimate errors. The other system (the *secondary* system) is composed of subsystems, no one of which is as effective as the primary system. The members of the secondary system may have poor detection capabilities, may detect targets only through infrequently exposed characteristics, or may possess large localization errors. They may however possess important characteristics which the primary system does not. For example the secondary system might be able to identify the target by nationality, by class (merchant or combatant), or by name. Although tracking unknown targets with the primary system may provide valuable information on target movements, the addition of information on target identity from a secondary source will increase the information content of both tracks and may provide essential data for analyzing tactics or predicting intentions.

It is assumed that throughout the major part of the surveillance operation the data from the two surveillance systems are processed independently and the report-to-track association and track continuation processes that were described are conducted separately for the two systems. Thus prior to the employment of the track-to-track association process the data base of the surveillance processing system contains tracks generated by the primary system and tracks generated by the secondary system. At certain times however the track families are examined to determine whether any of them refer to the same target and thus can be merged into a single track.

### BASIC MODEL

In this section the basic model employed in the investigation is described; this model involves one track from the primary system and one track from the secondary system. The primary system is denoted  $S_1$ , and its track is denoted  $T_1$ ; the secondary system is denoted  $S_0$ , and its track is denoted  $T_0$ .

Figure 1, which shows  $T_0$  and  $T_1$ , is representative of the basic problem formulation. The tracks have been generated by connecting target positions as generated by the two sensor systems. In the basic model each track consists of  $p$  positions, and the two systems are assumed to have both made their observations at the same times  $t_1, t_2, \dots, t_p$ . The  $i$ th position of track  $T_k$  is denoted by the coordinate pair  $(x_i^{(k)}, y_i^{(k)})$ , for  $i = 1, 2, \dots, p$  and  $k = 0, 1$ . Thus each  $T_k$  can be associated with the  $n$ -by-one (where  $n = 2p$ ) column vector

$$\begin{pmatrix} x_1^{(k)} \\ y_1^{(k)} \\ x_2^{(k)} \\ y_2^{(k)} \\ \dots \\ x_p^{(k)} \\ y_p^{(k)} \end{pmatrix}.$$

It is assumed that the n-tuples

$$(x_1^{(k)}, y_1^{(k)}, \dots, x_p^{(k)}, y_p^{(k)})$$

are observations from an n-dimensional normal distribution. For an arbitrary random variable  $x$ , let  $\bar{x}$  represent the expectation of  $x$ . Then each observed track  $T_k$  is regarded as a sample from an  $n = 2p$  dimensional normal distribution

$$\mathcal{N} \left[ \begin{pmatrix} \bar{x}_1^{(k)} \\ \bar{y}_1^{(k)} \\ \dots \\ \bar{x}_p^{(k)} \\ \bar{y}_p^{(k)} \end{pmatrix}; \Sigma_k \right],$$

where each  $\Sigma_k$  is a  $2p$ -by- $2p$  covariance matrix. The elements of the matrices  $\Sigma_k$  are determined by the characteristics of the sensor systems and possibly by the analytic methods used to obtain smoothed coordinates  $(x_i^{(k)}, y_i^{(k)})$  from the reported observed positions.

The test statistic to be used, denoted  $R^2$ , is defined by

$$R^2 = (T_0 - T_1)' \left( \sum_0 + \sum_1 \right)^{-1} (T_0 - T_1), \quad (1)$$

where the prime indicates matrix transpose. Given  $\Sigma_0$  and  $\Sigma_1$ , it is known [7] that there exists a nonsingular matrix  $P$  such that  $P(\Sigma_0 + \Sigma_1)P' = I$ , where  $I$  is the  $n$ -by- $n$  identity matrix. It will be shown that the statistic

$$W^2 = |P(T_0 - T_1)|^2 \quad (2)$$

has a distribution whose characteristics are useful in this investigation. Since it can be shown that  $W^2 = R^2$ , in the following we will consider that  $R^2 = |P(T_0 - T_1)|^2$  and use the notation  $r^2$  for observed values of  $R^2$  or for values which have been calculated from a specified set of track positions.

### NULL HYPOTHESIS $H_0$

The null hypothesis to be tested is  $H_0$ :  $T_0$  and  $T_1$  are from the same target. More specifically, let  $\bar{T}_k$  represent the column vector of expected values

$$\bar{T}_k = \begin{pmatrix} \bar{x}_1^{(k)} \\ \bar{y}_1^{(k)} \\ \dots \\ \bar{x}_p^{(k)} \\ \bar{y}_p^{(k)} \end{pmatrix}, \quad k = 0, 1.$$

Then an equivalent formulation of the null hypothesis is

$$H_0: \bar{T}_0 = \bar{T}_1$$

or

$$H_0: \bar{T}_0 - \bar{T}_1 = \mathcal{Q}$$

where  $\mathcal{Q}$  is the  $2p$ -by-1 column vector with all components equal to 0.

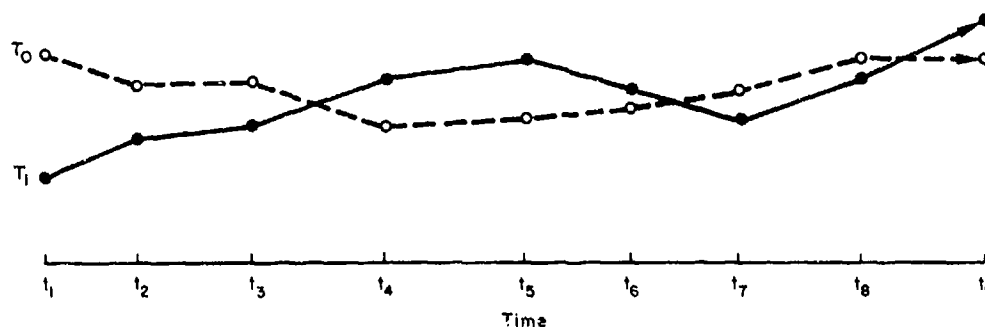


Fig. 1—The track association problem in the basic model, in which the primary system (subscript 1) and secondary system (subscript 0) observe track positions simultaneously

Since the tracks  $T_0$  and  $T_1$  are samples from  $n$ -dimensional normal distributions, with  $n = 2p$ , the difference  $T_0 - T_1$  is a sample from an  $n$ -dimensional normal distribution with expectation  $\bar{T}_0 - \bar{T}_1$  and covariance matrix  $\Sigma = \Sigma_0 + \Sigma_1$ . As mentioned, under  $H_0$ ,  $T_0 - T_1 = \mathbf{0}$ . With the matrix  $P$  defined as just before equation (2), it follows by properties of the multivariate normal distribution [8, pp. 24-25] that  $P(T_0 - T_1)$  is a sample from an  $n$ -dimensional normal distribution with zero mean and identity covariance matrix. Thus [7, p. 56] under  $H_0$ ,  $R^2$  has the (central) chi-square distribution with  $n = 2p$  degrees of freedom, with probability density function (p.d.f.)

$$f_0(r^2; n) = [2^p \Gamma(p)]^{p-1} e^{-r^2/2}. \quad (3)$$

The null hypothesis  $H_0$  is rejected whenever the computed value  $r^2$  of  $R^2$  seems too large. A type-I error occurs when the null hypothesis is incorrectly rejected; the probability of this error, the *level of significance*, is denoted  $\alpha$ . For a specified value of  $\alpha$  and a given number of observed positions, the acceptance region for  $H_0$  is bounded below by 0 and above by the value  $\chi^2$  satisfying the expression

$$1 - \alpha = \int_0^{\chi^2} f_0(r^2; n) dr^2. \quad (4)$$

Values of  $\chi^2$  for given  $\alpha$  and  $p$  are found in tables of the chi-square distribution, such as in Ref. 9; Table 1 contains representative values. For computer application either the chi-square tables can be stored in memory or analytic expressions can be used to compute approximations to  $\chi^2$ ; such approximations are discussed in Chapter 17 of Ref. 10 (Vol. 1).

It is valuable to consider the situation that inspires the rejection of the null hypothesis  $H_0$ ; these considerations lead to discussions of when the significance test ought to be carried out on two observed tracks. It is clear that the test should be carried out when and only when there is doubt as to the source of two observed tracks. As an example of a situation in which the test is *not* required, suppose there are tracks  $T_0$  and  $T_1$  as shown in Fig. 1, where the primary surveillance system has a high detection probability. For  $H_0$  to be rejected, the track  $T_0$  must be assumed to have come from a target other than that generating  $T_1$ . However, as shown in the figure, no such alternate track has been detected, and the existence of an alternate target is a low probability event in the light of the primary system's high detection probability. Thus, since no other target could have generated  $T_0$ , it is sufficient to associate the two tracks without recourse to the test.

It is not sensible to carry out the test unless rejection of  $H_0$  is operationally feasible. If the situation at hand does not admit the possibility of at least one additional target for association with  $T_0$ , then the two candidate tracks should be associated. Situations in which the test should be carried out will always admit the possibility of separate targets underlying the tracks  $T_0$  and  $T_1$ . For example, whenever the primary surveillance system has only a moderate detection probability, it is possible that a target may be present yet not be detected sufficiently often to have generated a track. Another possibility is that a number of candidate tracks have been detected by the primary system, so that should the test fail there will be other candidates for association with  $T_0$ . This case is discussed in the following section.

Table 1—Acceptance Regions for the Hypothesis  $H_0$ 

Number $p$ of Observed Positions in Each Track	Maximum Value $\chi^2$ for Acceptance of $H_0$				
	$\alpha = 0.30$	$\alpha = 0.20$	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$
2	4.878	5.989	7.779	9.488	13.277
4	9.524	11.030	13.362	15.507	20.090
6	14.011	15.812	18.549	21.026	26.217
8	18.418	20.465	23.542	26.296	32.000
10	22.775	25.038	28.412	31.410	37.566
15	33.530	36.250	40.256	43.773	50.892

ALTERNATE HYPOTHESIS  $H_1$ 

Although the track association test should be performed only when it is reasonable to suggest that track  $T_0$  could be related to a target other than the source of  $T_1$ , it is not always the case that a definite alternate source for  $T_0$  can be proposed. It is one thing to accept the possibility of an alternate source, but it is another to specify one. In the situation discussed in this section, an observed track is specified as relating to a possible alternate source for  $T_0$ . In this case one can specify an alternate hypothesis  $H_1$  and determine the power of the track association test against this alternative.

Figure 2, adapted from Ref. 11, illustrates an operational situation in which this case could arise. This figure contains examples of ship track histories a surveillance system such as the primary system might obtain over an interval of time for a given area of interest. Each of these tracks represents a separate target. This example contains both ships which appear to be moving randomly and ships which move in an apparently more predictable manner. The illustration shows much track crossing and few cases in which two or more tracks remain parallel for an extended period. Suppose that a track  $T_0$  from the secondary system appears to be related to one of the tracks in this illustration. If the association test should fail, then it would appear that there will be at most one other track with which  $T_0$  might be associated. If the primary system has a relatively high detection probability, then it is unlikely that it would not have detected the target which generated  $T_0$ . Consequently it is assumed that there will be another candidate track for  $T_0$  and thus that if association with the first track should fail, then association with the second track should be accepted.

Figure 3 illustrates the extension of the basic model to the case of testing the hypothesis  $H_0$  against a specific alternative. The tracks  $T_0$  and  $T_1$  are as before. Another track  $T_2$  is now present and is assumed to have been obtained from the primary system  $S_1$ . The track  $T_0$  must relate to one of the other tracks; if the hypothesis that  $T_0$  is associated with  $T_1$  is rejected, then  $T_0$  will be associated with  $T_2$ . The statistical distributions associated with  $T_0$  and  $T_1$  are as was described in the section on the basic model. The observations comprising  $T_2$  constitute a column vector

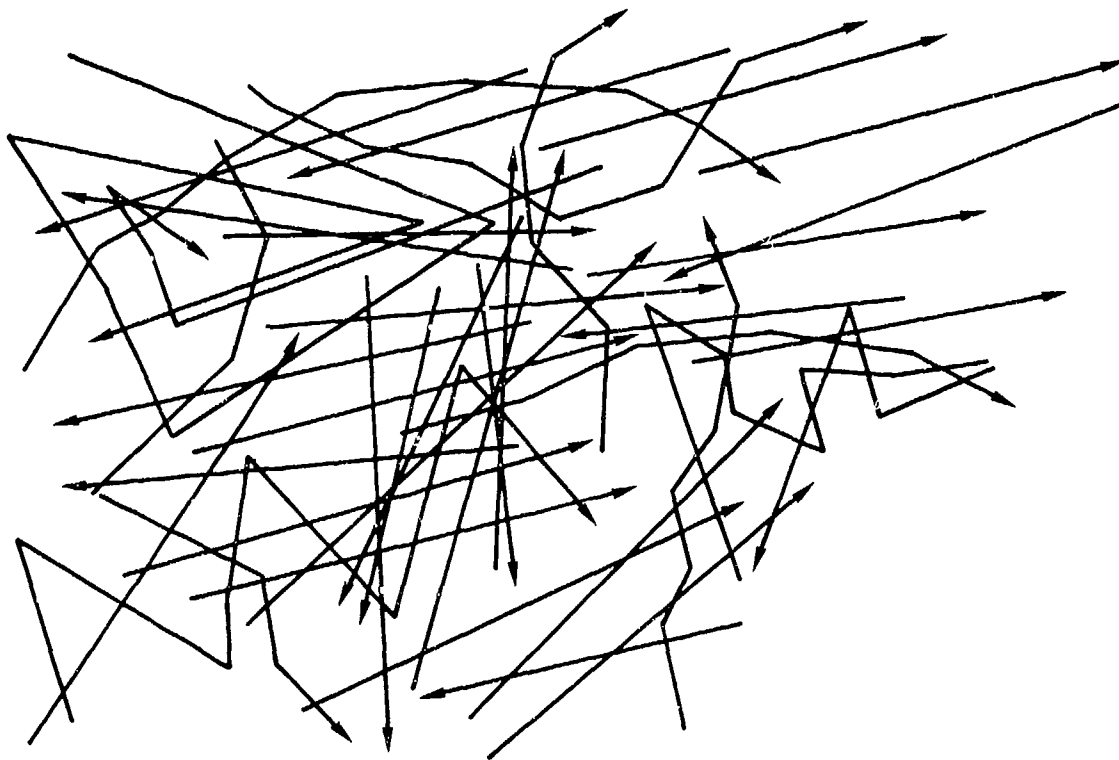


Fig. 2—Ship track histories from a single sensor.

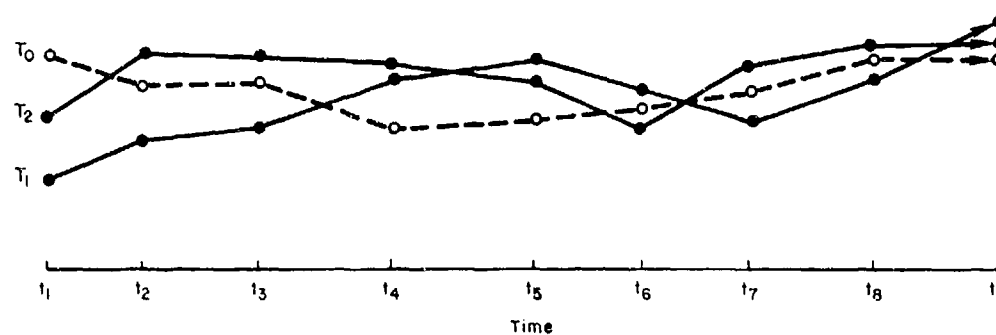


Fig. 3—Track association problem with two candidate tracks.

$$T_2 = \begin{pmatrix} x_1^{(2)} \\ y_1^{(2)} \\ x_2^{(2)} \\ y_2^{(2)} \\ \dots \\ x_p^{(2)} \\ y_p^{(2)} \end{pmatrix},$$

and this vector is assumed to be a sample from a  $2p$ -dimensional normal distribution with expectation

$$\bar{T}_2 = \begin{pmatrix} \bar{x}_1^{(2)} \\ \bar{y}_1^{(2)} \\ \bar{x}_2^{(2)} \\ \dots \\ \bar{x}_p^{(2)} \\ \bar{y}_p^{(2)} \end{pmatrix}$$

and  $2p$ -by- $2p$  covariance matrix  $\Sigma_2$ . It is clear that this case admits a specified alternate hypothesis  $H_1$ , which states that  $T_0$  is related to  $T_2$ . The alternate is given by

$$H_1: T_0 = T_2.$$

#### Distribution of $R^2$ Under $H_1$ ; Power of the Test

Under  $H_1$  the difference vector  $T_0 - T_1$  is a sample from a  $2p$ -dimensional normal distribution with expectation  $\bar{T}_2 - \bar{T}_1$  and covariance matrix  $\Sigma = \Sigma_0 + \Sigma_1$ . Thus [7, p. 56]  $R^2$  has the *noncentral chi-square distribution* with  $n = 2p$  degrees of freedom and noncentrality parameter

$$\begin{aligned} \lambda &= (\bar{T}_2 - \bar{T}_1)' \Sigma^{-1} (\bar{T}_2 - \bar{T}_1) \\ &= |P(\bar{T}_2 - \bar{T}_1)|^2, \end{aligned}$$

where matrix  $P$  is as defined previously. The probability density function of the noncentral chi-square distribution with  $n$  degrees of freedom and noncentrality  $\lambda$  is defined by

$$f_1(r^2; n, \lambda) = \frac{1}{2} \left( \frac{r^2}{\lambda} \right)^{(n-2)/4} e^{-(r^2+\lambda)/2} I_{(n-2)/2}(r\sqrt{\lambda}), \quad r \geq 0,$$

where  $I_K(\dots)$  is the modified Bessel function of the first kind and order  $K$  [Ref. 10]. Thus for a given set of  $p$  positions, a specified value for the level of significance  $\alpha$ , and specified value of  $\lambda$ , the probability of incorrectly accepting the hypothesis  $H_0$  when  $H_1$  is true is

$$\beta = \int_0^{\chi^2} f_1(r^2; n, \lambda) dr^2, \quad (5)$$

where  $n = 2p$  and where the upper limit  $\chi^2$  is defined by expression (4). The power of the test,  $\pi = 1 - \beta$ , is given by

$$\pi = \int_{\chi^2}^{\infty} f_1(r^2; n, \lambda) dr^2. \quad (6)$$

For track-to-track association the test of  $H_0$  against  $H_1$  is a test of the central chi-square distribution against a noncentral chi-square alternative. Tables relating to this test are available [for example, Ref. 12]. Reference 13 describes a rapid, compact program for carrying out computations based on the noncentral chi-square distribution. This program is suitable for implementation within automated tracking algorithms or for use within computer-driven analyst aids.

Figure 4 indicates the power of the test for levels of significance of 0.10 and 0.01 and for situations involving six and 15 observed positions in each track. When all other parameters are held fixed, the power of the test increases if  $\lambda$  increases, increases if the specified level of significance increases, and decreases if the number of observed positions increases. Generally speaking, the operational situation will determine the values of  $p$ ,  $T_0$ ,  $T_1$ ,  $T_2$ ,  $\Sigma_0$ , and  $\Sigma_1$ . The error probabilities  $\alpha$  and  $\beta$  can be specified. However, the value of  $\lambda$  must be estimated by the surveillance analyst, since the value of the difference vector  $\bar{T}_2 - \bar{T}_1$  cannot be assumed to be known. An important question underlying the test procedure is the sensitivity of the resulting errors  $\alpha$  and  $\beta$  to variations in the  $\lambda$  value employed. It must be determined whether the test procedure requires great accuracy on the part of the analyst.

#### Sensitivity to Estimates of $\lambda$

The parameter  $\lambda$  is determined by the difference vector  $\bar{T}_2 - \bar{T}_1$  and by the matrix sum  $\Sigma = \Sigma_0 + \Sigma_1$ ; since the latter factor will generally be known, the requirement that  $\lambda$  be known reduces to a requirement that the difference vector be estimated, since it cannot be known exactly. The effects of variation in either the difference vector or in  $\lambda$  on the results of the test procedure can be ascertained only by a detailed parametric

analysis of the relations between  $p$ ,  $\alpha$ ,  $\beta$ ,  $\bar{T}_2 - \bar{T}_1$ , and  $\Sigma$ . Some insights into the general problem of test sensitivity to  $\lambda$  can be obtained from an analysis of the two-candidate test in the following special case. In the two-candidate track association problem the choice of which track to use as a basis for the null hypothesis  $H_0$  is arbitrary. Therefore in the absence of any supplementary information it will be assumed that errors in failing to make a correct association are as important for the alternate hypothesis as for the null hypothesis. In this case it is reasonable to set  $\alpha = \beta$ . Having specified  $\alpha$ , one can proceed as before by finding that value of  $\chi^2$  such that equation (4) holds. Using this value as the test criterion, one can then employ equation (5) to determine that value of  $\lambda$  such that the resulting  $\beta$  value is equal to  $\alpha$ . Analysis of the resulting relations between  $\alpha (= \beta)$  and  $\lambda$  can indicate how precisely  $\lambda$  must be estimated to obtain a test with level of significance  $\alpha$  and power  $1 - \alpha$ .

Figure 5 shows the sensitivity of the resulting test errors to estimates of  $\lambda$  by exhibiting contours along which  $\alpha = \beta$  for  $p$  values of 6, 15, and 24. Since the power of the test is an increasing function of  $\lambda$ , this figure shows that to guarantee a specified maximum error probability it is necessary only to guarantee that the actual value of  $\lambda$  exceeds a stated threshold. For example, for the case of 15 observed track positions, if one can guarantee that  $\lambda$  is greater than 30, then the test will achieve error rates of less than 8 percent. The lower threshold for estimates of  $\lambda$  increases as the number of observed positions increases and as the allowable maximum error probability decreases. Therefore, although all that the procedure requires is a decision that  $\lambda$  exceed some stated minimum value, it may be difficult to assure that this decision is correct.

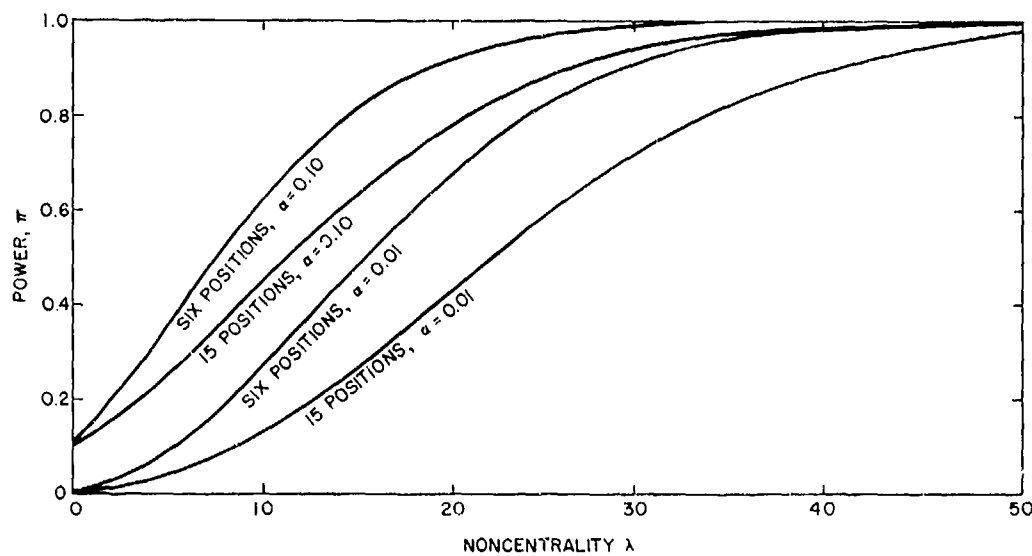


Fig. 4—Power of the test of  $H_0$  against  $H_1$ .

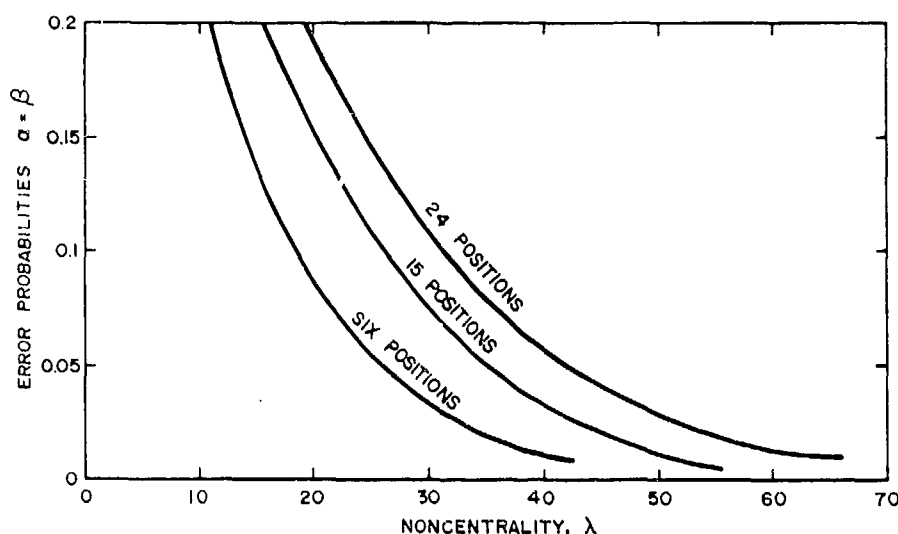


Fig. 5—Sensitivity of the test of  $H_0$  against  $H_1$  to changes in the noncentrality, for  $\alpha = \beta$ .

An example will illustrate the concepts involved in relating requirements on  $\lambda$  to requirements on the estimated value of  $\bar{T}_2 - \bar{T}_1$ . Suppose that tracks  $T_1$  and  $T_2$  were observed with a sensor which generated a covariance matrix with no off-diagonal elements and such that both  $\sigma_x^2$  and  $\sigma_y^2$  equaled 1 n.mi.<sup>2</sup> for each observed position. Suppose that  $T_0$  was generated by a system whose covariance matrix also was diagonal, but with  $\sigma_x^2 = \sigma_y^2 = 4$  n.mi.<sup>2</sup> for each position. Finally, suppose six positions (12 coordinates) have been observed and that an analyst can estimate that each element in the difference vector  $\bar{T}_2 - \bar{T}_1$  is at least 2.0 n.mi. Then the minimum value for  $\lambda$  will be 9.6, and one can conclude from Fig. 5 that the test could be subject to errors greater than 20 percent. Suppose on the other hand that errors of at most 5 percent were required for tracks comprising 15 observed positions. This would require a value of  $\lambda$  of at least 35, and with the covariance matrix structure described it would be necessary to be correct in deciding that the elements in the difference vector were, on the average, at least 2.4 n.mi.

### Computer Implementation

Computer implementation of a procedure based on these results could be readily achieved as an interactive analyst aid. Since the majority of the required parameter values would already be available, an analyst would have to enter his best judgment of the magnitude of the elements in the underlying mean difference vector. The program would then generate the test procedure, identifying the operable value of  $\alpha$ , computing the value of the upper limit  $\chi^2$ , computing the value of  $r^2$ , and determining whether or not  $r^2$  exceeded  $\chi^2$ . It would respond to the analyst with a statement of the decision which was made, together with an indication of the error probabilities present in the test.

used. Conversely, if an analyst were to specify a desired level of significance  $\alpha$ , the program could respond with the minimum value of  $\lambda$  required to guarantee an error  $\beta$  no greater than  $\alpha$ .

## FURTHER ASPECTS OF THE PROBLEM

In this section the preceeding discussion is extended to three additional topics. The first is a method for treating the situation when the basic assumption of simultaneous observations is relaxed. The second is the case in which immediate association decisions are not always required and can be deferred pending the receipt of further data. Finally the third is the importance of selecting proper limits for the error probabilities  $\alpha$  and  $\beta$  in the tests and the consequences of making track association errors within a surveillance-data processing system.

### Nonsimultaneous Observations

The basic track-to-track association model assumed that the positions constituting the two tracks were observed at the same times. Usually different surveillance systems will make detections at different times; hence for any pair of generated tracks a set of nonsimultaneous observations is more likely than a set of simultaneous observations. The following discussion is a description of a method of testing tracks with nonsimultaneous observations for equality of means. The method transforms the case of nonsimultaneous observations to the case of simultaneous observations. This transformation involves the generation of interpolated points within each track so that both of the resulting new tracks will have constituent points associated with the same set of observation times. The set of observation times contains all times at which either of the two original tracks was observed.

Figures 6 and 7 represent the situation. Figure 6 shows track  $T_1$ , observed at times  $t_2, t_3, t_4$ , and  $t_6$ , and track  $T_0$ , observed at times  $t_1, t_5$ , and  $t_7$ . To test the hypothesis that these tracks relate to the same target, two new tracks  $T_1^*$  and  $T_0^*$  have been created (Fig. 7), both of which comprise all the observation times for both  $T_1$  and  $T_0$ . The points in the new tracks are of two types. First, all points originally observed for a track are contained in the related new track. Second, if an original track  $T_k$  contains two sequential points, observed at times  $t_i$  and  $t_j$ , and if the other track contains a point observed at an intermediate time  $t$ , where  $t_i < t < t_j$ , then an intermediate point is interpolated between the original points. Letting  $D = (t - t_i)/(t_j - t_i)$ , the new point is assumed to occur at time  $t$  and to be located at a fraction  $D$  of the way between the originally observed points. If  $x_i$  and  $x_j$  are the  $x$  coordinates of  $T_k$  at times  $t_i$  and  $t_j$  respectively and  $x^*$  is the  $x$  coordinate of the new interpolated point in track  $T_k^*$  at time  $t$ , then

$$\begin{aligned} x^* &= x_i + D(x_j - x_i) \\ &= \frac{(t_j - t)}{(t_j - t_i)} x_i + \frac{(t - t_i)}{(t_j - t_i)} x_j. \end{aligned}$$

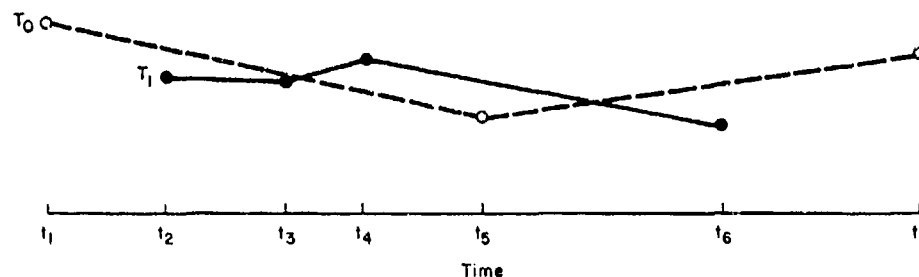


Fig. 6—Tracks with Nonsimultaneous observations

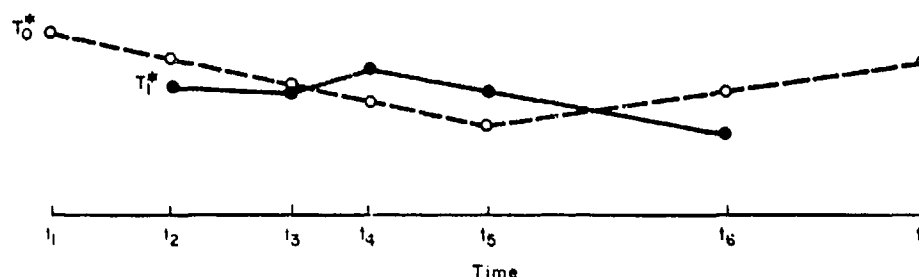


Fig. 7—Nonsimultaneous tracks of Fig. 6 after transformation to the case of simultaneous observations

These three points, the two original ones and the new interpolated one, are all made members of the new track  $T_k^*$ . The process is carried out for both of the original tracks and all of the time points for which the interpolation is feasible, resulting in the situation shown in Fig. 7. In this figure the two new tracks  $T_1^*$  and  $T_0^*$  both have "observation points" at all times  $t_2$  through  $t_6$ . Since these points have been obtained by a linear transformation of normal random variables, their related covariance structures can be obtained [8] and they may be used as data for the test of the null hypothesis  $H_0$  described for the basic model. The one-point extensions in track  $T_0^*$ , both at the start and at the end, are not used in the hypothesis-testing computations, as the interpolation scheme does not generate extensions from track  $T_1$ .

This extension of the basic model to the case of nonsimultaneous observations has been programmed for use within an automated track correlation process. This program is described in Appendix A and listed in Appendix B.

### Sequential Tests

A situation was discussed in which track  $T_0$  was to associate with one of two candidate tracks  $T_1$  and  $T_2$  as a result of a testing procedure. It is not always the case that an immediate track association decision is required; quite often the problems inherent in making association errors are severe enough to require that association decisions not be made without sufficient supporting evidence. One method of accomplishing this is through the use of sequential tests based on reexamination of the data set each time new

data elements are obtained. In the following discussion the structure of sequential testing procedures within the context of the track association problem is described and some problems inherent in implementing the procedure are indicated.

As described in Refs. 14 and 15, sequential testing procedures are useful in deciding whether observed data correspond to a null hypothesis  $H_0$ , whether they correspond to a specified alternate hypothesis  $H_1$ , or whether additional data should be obtained before making a final decision. Let  $g_{0,n}$  and  $g_{1,n}$  be the probability density functions of the multivariate data set  $(x_1, x_2, \dots, x_n)$  at the  $n$ th step of the process, under hypothesis  $H_0$  and  $H_1$  respectively, let

$$h_n = \frac{g_{1,n}(x_1, x_2, \dots, x_n)}{g_{0,n}(x_1, x_2, \dots, x_n)},$$

and suppose that the error probabilities  $\alpha$  and  $\beta$  have been specified. The sequential testing procedure is then defined by the following criteria:

$$\text{if } h_n \leq \frac{\beta}{1-\alpha}, \quad \text{accept } H_0; \quad (7a)$$

$$\text{if } \frac{\beta}{1-\alpha} < h_n < \frac{1-\beta}{\alpha}, \quad \text{obtain another observation} \quad (7b)$$

$$\text{if } \frac{1-\beta}{\alpha} \leq h_n, \quad \text{accept } H_1. \quad (7c)$$

Note that the criteria do not depend on the sample size  $n$ . These criteria were based on the assumptions that the successive observations were stochastically independent samples and that the sequential procedure will, with probability 1, eventually terminate. For a rich family of situations the assumption of stochastically independent samples can be relaxed and the procedure will still be valid. It has been assumed here that the conditions underlying the track association problem are such that the sequential test procedure is valid.

For track association the assumptions of underlying normal distributions imply that under the null hypothesis  $H_0$  specified earlier  $T_0 - T_1$  has the density function

$$\begin{aligned} g_{0,n} &= g_0(T_0 - T_1; p) \\ &= \frac{1}{(2\pi)^p |\Sigma|^{1/2}} \exp \left[ -\frac{1}{2} (T_0 - T_1)' \Sigma^{-1} (T_0 - T_1) \right] \end{aligned}$$

and under  $H_1$  specified earlier  $T_0 - T_1$  has the density function

$$g_{1,n} = g_1(T_0 - T_1; p, \bar{T}_2 - \bar{T}_1) \\ = \frac{1}{(2\pi)^p |\Sigma|^{1/2}} \exp \left( -\frac{1}{2} \{[(T_0 - T_1) - (\bar{T}_2 - \bar{T}_1)]' \Sigma^{-1} [(T_0 - T_1) - (\bar{T}_2 - \bar{T}_1)]\} \right).$$

It follows that

$$h_n = \frac{g_{1,n}}{g_{0,p}} \\ = \exp \left\{ -\frac{1}{2} [-(T_0 - T_1)' \Sigma^{-1} (\bar{T}_2 - \bar{T}_1) - (\bar{T}_2 - \bar{T}_1)' \Sigma^{-1} (T_0 - T_1) \right. \\ \left. + (\bar{T}_2 - \bar{T}_1)' \Sigma^{-1} (\bar{T}_2 - \bar{T}_1)] \right\} \\ = \exp \left\{ -\frac{1}{2} [\lambda - 2(T_0 - T_1)' \Sigma^{-1} (\bar{T}_2 - \bar{T}_1)] \right\}.$$

As discussed previously, it is reasonable to consider that the error probabilities  $\alpha$  and  $\beta$  are equal. Thus, with the notation

$$Z = \lambda - 2(T_0 - T_1)' \Sigma^{-1} (\bar{T}_2 - \bar{T}_1),$$

criteria (7) reduce to the following:

$$\text{if } \ln \left( \frac{1-\alpha}{\alpha} \right)^2 \leq Z, \quad \text{accept } H_0, \quad (8a)$$

$$\text{if } \ln \left( \frac{\alpha}{1-\alpha} \right)^2 < Z < \ln \left( \frac{1-\alpha}{\alpha} \right)^2, \quad \text{wait for additional data,} \quad (8b)$$

$$\text{if } Z \leq \ln \left( \frac{\alpha}{1-\alpha} \right)^2, \quad \text{accept } H_1. \quad (8c)$$

The performance of this procedure is determined by the relations between the observed values  $T_0 - T_1$  and  $\Sigma$  and the estimated mean difference vector  $\bar{T}_2 - \bar{T}_1$ . Detailed investigations based on realistic predictions of the observed parameter values will be required to ascertain the sensitivity of the sequential procedure to estimates of the difference vector in an operational setting. At this point however examination of the criteria (8) permit general conclusions to be made regarding the test's performance. Figure 8 illustrates these criteria. If the permissible error probabilities are low, the hypothesis acceptance regions are reduced; the sequential procedure will most frequently decide to wait for additional data unless a value of  $Z$  is obtained which is large in absolute value. If the test is more lenient—reflected in higher allowable error probabilities—the sequential procedure will more frequently result in a hypothesis acceptance, unless the absolute value of  $Z$  is small.

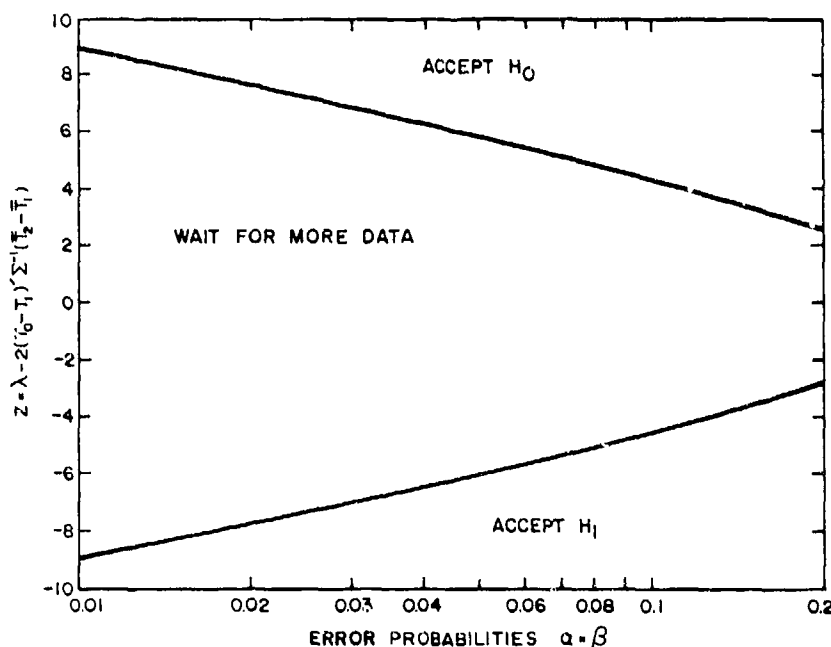


Fig. 8—Sequential test criteria (8) for track association.

In summary, the use of sequential testing procedures may be valuable when two tracks  $T_1$  and  $T_2$  are candidates for association. This test permits additional information to be collected if the available information is adequate for a hypothesis acceptance decision. The test procedure requires the estimation of the mean difference vector  $\bar{T}_2 - \bar{T}_1$ . The sensitivity of the test's performance to estimates of this value is a function of the observed data  $T_0 - T_1$  and covariance matrix  $\Sigma$ ; further detailed investigations of this factor will be required to determine the worth and feasibility of implementing the sequential procedure within an operational surveillance processing system.

#### Specifying the Allowable Error Probabilities

The tests described in this report require the specification of maximum allowable probabilities  $\alpha$  and  $\beta$  of incorrect decisions. These errors have real-world costs. Those responsible for implementing these tests in an operational surveillance-data processing system must analyze the consequences of making incorrect decisions in order to specify  $\alpha$  and  $\beta$ .

In the case of the basic model the only decision is whether or not to associate a given pair of tracks; as developed, the test requires the specification of  $\alpha$ , the probability of incorrectly rejecting the proposed association of the two tracks. Errors of this type will require the surveillance processing system to maintain multiple tracks in situations when maintenance of single tracks would suffice. The data base will thus become cluttered with redundant tracks, and any computational routines that are performed on each track in the data base will be performed more than necessary. More important is the possibility of losing information on special-interest targets which correct track association

## HOWARD L. WIENER

would have provided. For example, without supporting information on target identities, position-only tracks must all be subject to the same degree of attention by analysts. However, if position-only tracks can be associated with tracks carrying identification information, then known combatant tracks can be specified for special attention and tracks from noncombatants can be specified for lesser attention. Thus a low false rejection rate for  $H_0$  will increase data-processing efficiency and increase the value of the information in the stored tracks. Still the consequences of falsely associating unrelated tracks can be severe. Whereas false rejection results in the loss of useful information, false acceptance surely results in the generation of misleading information. Misclassification of a combatant as a merchant, or conversely, can certainly create false pictures of tactical situations. Track files containing incorrectly associated tracks can easily serve as bases for future compounded errors, thus leading to a possibly useless data base. Extensive lists of the consequences of making either type of error can be generated indefinitely, but these examples should suffice to indicate the importance of proper assessment of the effects of the errors on the operation and ultimate value of the data processing system.

## SUMMARY REMARKS

This report has discussed a series of statistical tests for attacking the track-to-track association problem in ocean surveillance. The tests are straightforward and can be readily implemented within many automated surveillance-data processing systems. Computational algorithms exist which permit rapid computation of any of the test statistics involved; consequently these test procedures can be implemented "on line," within an automatic processing system, or they can be used within interactive programs to provide quick response to analyst queries. The demand for inputs from analysts will be at a low level; optimally, analysts may be asked simply to judge whether a given parameter is or is not within a specified range. A major value of these tests lies in their permitting greater use of the various types of information contained in different tracks of the same target. The result of successful implementation of the tests is a data base which is low in errors and low in redundancy.

## ACKNOWLEDGMENTS

The author expresses his appreciation to David J. Kaplan for providing numerous helpful suggestions on the report's exposition and to Dr. Joseph H. Kullback for suggesting the problem and providing insights into practical aspects of surveillance-data processing systems.

## REFERENCES

1. M.E.B. Owens, "A Probability Density Function for the Future Position of a Vessel at Sea," NRL Report 7128, Aug. 1970.
2. M.E.B. Owens, "Probabilities of Successful Data Association for Tracking in the Plane when Using Maximum Likelihood Association Procedures," NRL Report 7162, Feb. 1971.

NRL REPORT 7989

3. H.L. Stalford, "Solitary Occupancy for Unequal Cell Probabilities with Application to Doppler Radars for Ocean Surveillance," NRL Report 7708, May 1974.
4. H.L. Stalford, "A Minimum-Distance Method for Ocean-Surveillance Resolution," NRL Report 7677, May 1974.
5. M.E.B. Owens and J.H. Kullback, "Two-Dimensional Point Resolution," NRL Report 7122, Sept. 1970.
6. F.K. Bechtel and D.E. Amos, "Difference Detection, Tradeoffs Between Detection Probability and Processing Complexity," IEEE Transactions on Systems, Man, and Cybernetics SMC-5 (No. 4), 446-452 (July 1975).
7. K.S. Miller, *Multidimensional Gaussian Distributions*, New York, Wiley, 1964.
8. T.W. Anderson, *An Introduction to Multivariate Statistical Analysis*, New York, Wiley, 1958.
9. R.A. Fisher and F. Yates, *Statistical Table for Biological, Agricultural, and Medical Research*, New York, Hafner, 1948.
10. N.L. Johnson and S. Kotz, *Distributions in Statistics, Continuous Univariate Distributions*, Vols. 1 and 2, Boston, Houghton Mifflin, 1970.
11. IBM Federal Systems Division Report 16-955-3923 (Secret), 29 June 1973.
12. G.E. Haynam, Z. Govindarajulu, and F.C. Leone, "Tables of the Cumulative Non-Central Chi-Square Distribution," in *Selected Tables in Mathematical Statistics*, Vol. 1, H.L. Harter and D.B. Owen, editors, Chicago, Markham, 1970.
13. H.L. Wiener, "A Fortran Program for Rapid Computations Involving the Non-Central Chi-Square Distribution," NRL Memorandum Report 3106, Aug. 1975.
14. A. Wald, *Sequential Analysis*, New York, Wiley, 1947.
15. B.W. Lindgren, *Statistical Theory*, New York, Macmillan, 1960 (2nd edition, 1968).
16. M. Abramowitz and I.A. Stegun, editors, *Handbook of Mathematical Functions*, National Bureau of Standards, (Washington, D.C., U.S. Govt. Printing Office) 1964.

## Appendix A

### A TRACK-TO-TRACK ASSOCIATION PROGRAM

#### INTRODUCTION

A program for track-to-track association tests is described in this appendix and listed in Appendix B. The program is designed to operate on tracks which are defined by sets of positions and which are accompanied by the underlying covariance matrices. The track positions may comprise locations reported by a surveillance sensor or may consist of estimated positions obtained by applying a smoothing scheme to a set of observations. In the former case the covariance matrix would probably be block diagonal, with two-by-two matrices along the main diagonal; in the latter case the covariance matrix would be a general nonsingular symmetric matrix, possibly with relatively large off-diagonal entries. In any event the multivariate covariance structure underlying the positions constituting the tracks must be known and is to be entered as data. The scheme is based on the assumption that the degree of overhang for overlapping tracks is limited to one point at each end of the tracks; that is, when two tracks start at different times, the early track contains at most one early point and similarly a track which continues after the other has stopped contains at most one late point. Finally the scheme will test track associations for track positions observed either at identical times or at different times.

#### PROGRAM STRUCTURE

The following steps comprise the major activities of the track association program.

*Read Inputs.* For each point in each track the program first obtains the time of observation, the  $X$  coordinate, and the  $Y$  coordinate. In addition, it obtains the covariance matrix related to the set of points constituting each track. Finally it generates a column vector  $TRK$  for each observed track, defined such that for the  $K$ th observed track ( $K = 1, 2$ ) the  $I$ th observed ( $X, Y$ ) coordinate pair is defined as the pair

$$(TRK(K, 2I - 1), TRK(K, 2I)).$$

The times  $T$  used in the program are assumed to be defined in terms of decimal hours; some changes will be required whenever actual inputs are given in terms of hours, minutes, and seconds.

*Compute the Interpolation Parameters.* The fundamental concept of the program is the coordinate-by-coordinate comparison of a pair of standardized interpolated tracks. The program obtains these from the track vectors  $TRK(1)$  and  $TRK(2)$  by a linear interpolation scheme which produces a set of identical time points and related positions within each track. The interpolation scheme proceeds as described in the main body of this report in conjunction with Fig. 6 and 7. Next, if the starting time for one interpolated track is earlier than

that for the other one, the program deletes the earlier one, so that both interpolated tracks will now start at the same time; similarly any overhang at the termination of the tracks is reduced to the case of equal track termination time. Thus both interpolated tracks will be based on the same set of observed time points. The number of such points is denoted NPTS; since each observation point is related to two coordinates, the majority of the subsequent computation is based on twice the value of this parameter.

*Compute the Transformation Matrices for the Original Tracks.* The interpolated tracks used for the statistical test are obtained from the original tracks by linear transformations. If the input track is based on  $M$  observed points ( $2M$  coordinates), the transformation matrix  $Q$  has 2 ( $NPTS$ ) rows and  $2M$  columns.

*Compute the Interpolated Tracks.* For each  $K = 1, 2$ , input track  $TRK(K)$  is transformed into interpolated track  $TINT(K)$  via the transformation matrix  $Q$ , with suitable dimensions. Since the input track is treated as a column vector, the equation for this transformation is

$$TINT = Q \cdot TRK.$$

*Compute the Covariance Matrices for the Interpolated Tracks.* The covariance matrix  $\Sigma_{TI}$  underlying the related interpolated track is given by

$$\Sigma_{TI} = Q \Sigma_T Q',$$

where  $\Sigma_T$  is the covariance matrix for an originally obtained track and the prime represents matrix transpose.

*Compute the Difference Vector for Interpolated Tracks.* The hypothesis that the two original tracks are from the same target is equivalent to the hypothesis that the underlying mean vectors are equal, or that the difference of the mean vectors contains all zero coordinates. This formulation carries over to the interpolated vectors as well, and at this step the program computes the value of the difference  $TDIFF$  of the computed interpolated tracks,  $TDIFF$  being a column vector.

*Compute the Matrix Sum.* In addition to the values of the difference vector, the test requires knowledge of the covariance matrix associated with the difference, in this case the sum  $\Sigma$  of the covariance matrices  $\Sigma_{TI}$  associated with the interpolated tracks.

*Compute the Measure  $R^2$ .*  $R^2$  is calculated by the formula

$$R^2 = (TDIFF)' \Sigma^{-1} (TDIFF),$$

where again the prime indicates matrix transpose.

*Compute Allowable Upper Limit for  $R^2$ .* The program next computes the upper limit  $\chi^2$ , above which computed values of  $R^2$  will lead to rejection of the hypothesis that the two tracks are from the same target. This limit is the solution of expression

(4) of the main body of the report. These computations use two approximations. The first one, based on equation (28) in Chapter 17 of Ref. 10 (Vol. 1), permits approximation of the chi-square cumulative distribution function (c.d.f.) by the distribution function of an appropriately defined normal distribution. The second one, based on expression 26.2.23 of Ref. 16, uses the ratio of two polynomials to estimate percentiles of normal distributions. This method requires relatively little execution time and produces results which are within a few percentage points of values in the standard statistical tables.

*Make Decision.* The program compares the computed value of  $R^2$  with the upper limit  $\chi^2$ . If  $R^2 \leq \chi^2$ , the program sends a message to merge the two tracks. Otherwise, it sends a message not to merge the tracks based on the available data.

## PROGRAM LISTING

Appendix B is a listing of the program, as written for implementation on the CDC KRONOS time-sharing computer system. In the KRONOS version the array sizes reflect the small number of data points used for the test runs; for actual data the array sizes may have to be increased. Also, the following items are specific for KRONOS and will have to be changed for other modes of operation, such as batch processing or implementation within an automated correlation processor:

- The array W1 and W2;
- The CALL MATIDN instruction (line 00160);
- The instruction  $W2(I, J) = S1GM(I, J)$  (line 03320);
- The CALL MATINV instruction (line 03660).

## ADDITIONAL PROGRAMS

An extended version of this program, which uses Kalman filtering techniques to obtain smoothed position estimates and the related covariance matrix, has been developed. Information on the extended program and on other programs related to the track association problem can be obtained from the author.

Appendix B  
PROGRAM LISTING

LIST

75/03/20. 11.28.24.  
PROGRAM ASSDC

```
00020 PROGRAM ASSDC(OUTPUT,TAPE1)
00040 DIMENSION DUM(16,16), M(2), NUM(2), Q(2,16,16), SG(2,8,8)
00060 DIMENSION SIG(2,16,16), SIGI(16,16), SIGM(16,16), T(8)
00080 DIMENSION TCDM(8), TDIFF(16), TINT(2,16), TM(2,8,8)
00100 DIMENSION TRK(2,8), TS(2,8), X(2,4), Y(2,4)
00120 DIMENSION W1(16,32), W2(16,16)
00140 ALPHA=0.05
00160 CALL MATIDN(W2,16,16)
00180*
00200* INPUT DATA
00220 READ(1,5050) M(1)
00240 5050 FORMAT(I3)
00260 M2=M(1)
00280 M22=M2+M2
00300 DO 20 I=1,M2
00320 READ(1,5070) TS(1,I),X(1,I),Y(1,I)
00340 5070 FORMAT(6F8.4)
00360 20 CONTINUE
00380 DO 30 I=1,M22
00400 READ(1,5070) (SG(1,I,J),J=1,M22)
00420 30 CONTINUE
00440 DO 50 I=2,M22
00460 I1=I-1
00480 DO 40 J=1,I1
00500 SG(1,I,J)=SG(1,J,I)
00520 40 CONTINUE
00540 50 CONTINUE
00560 READ(1,5050) M(2)
00580 M2=M(2)
00600 M22=M2+M2
00620 DO 60 I=1,M2
00640 READ(1,5070) TS(2,I),X(2,I),Y(2,I)
00660 60 CONTINUE
00680 DO 70 I=1,M22
00700 READ(1,5070) (SG(2,I,J),J=1,M22)
00720 70 CONTINUE
00740 DO 90 I=2,M22
00760 I1=I-1
00780 DO 80 J=1,I1
00800 SG(2,I,J)=SG(2,J,I)
00820 80 CONTINUE
00840 90 CONTINUE
00860 DO 100 K=1,2
00880 M2=M(K)
00900 DO 95 I=1,M2
00920 TRK(K,I+1-I)=X(K,I)
00940 TRK(K,I+I)=Y(K,I)
00960 95 CONTINUE
00980 100 CONTINUE
```

HOWARD L. WIENER

```

01000 PRINT 9000
01020 9000 FORMAT(30X,2H1.INPUTS)
01040 DO 140 K=1,2
01060 M2=M(K)
01080 M22=M2+M2
01100 PRINT 9010, K
01120 9010 FORMAT(//,3HTRACK NO.,I2/3X,4HTIME,7X,7HX-COORD,
01140+5X,7HY-COORD)
01160 DO 130 I=1,M2
01180 PRINT 9020, TS(K,I), X(K,I), Y(K,I)
01200 9020 FORMAT(5(F10.4, 2X))
01220 130 CONTINUE
01240 PRINT 9030,K
01260 9030 FORMAT(//,20X,21HCOVARIANCE MATRIX NO., I2)
01280 DO 135 I=1,M22
01300 PRINT 9040, (SG(K,I,J), J=1,M22)
01320 9040 FORMAT(6(F10.4,1X))
01340 135 CONTINUE
01360 140 CONTINUE
01380*
01400* COMPUTE THE INTERPOLATION PARAMETERS
01420 CALL ORDER(TS,M,T,N)
01440 DO 160 K=1,2
01460 M1=M(K)-1
01480 NUM(K)=0
01500 DO 155 I=1,M
01520 DO 150 J=1,M1
01540 IF ((T(I).GE.TS(K,J)).AND.(T(I).LE.TS(K,J+1))) GO TO 145
01560 GO TO 150
01580 145 TM(K,I,J)=(TS(K,J+1)-T(I))/(TS(K,J+1)-TS(K,J))
01600 TM(K,I,J+1)=(T(I)-TS(K,J))/(TS(K,J+1)-TS(K,J))
01620 NUM(K)=NUM(K)+1
01640 GO TO 155
01660 150 CONTINUE
01680 155 CONTINUE
01700 160 CONTINUE
01720 IF(TS(1,1).EQ.TS(2,1)) GO TO 320
01740 IF(T(1).EQ.TS(1,1)) GO TO 200
01760 I=2
01780 GO TO 240
01800 200 I=1
01820 GO TO 240
01840 240 K=N-1
01860 NUM(I)=NUM(I)-1
01880 DO 300 J=1,K
01900 DO 280 I=1,2
01920 MX=M(I)
01940 DO 260 L=1,MX
01960 TM(I,J,L)=TM(I,J+1,L)
01980 260 CONTINUE
02000 280 CONTINUE
02020 TCOM(J)=T(J+1)
02040 300 CONTINUE
02060 NPTS=MIN0(NUM(1),NUM(2))
02080 GO TO 400
02100 320 NPTS=MIN0(NUM(1),NUM(2))
02120 DO 360 J=1,NPTS
02140 TCOM(J)=T(J)
02160 360 CONTINUE
02180 400 CONTINUE

```

# NRL REPORT 7989

```

02200*
02220*   CONSTRUCTION OF THE MATRICES Q1,Q2 FOR TRANSFORMING
02240*   TRACKS 1 AND 2, RESPECTIVELY.
02260 DO 740 K=1,2
02280 M2=M(K)
02300 DO 720 I=1,NPTS
02320 DO 700 J=1,M2
02340 INEW=I+I-1
02360 JNEW=J+J-1
02380 Q(K, INEW, JNEW)=TM(K, I, J)
02400 INEW=I+I
02420 JNEW=J+J
02440 Q(K, INEW, JNEW)=TM(K, I, J)
02460 700 CONTINUE
02480 720 CONTINUE
02500 740 CONTINUE
02520*
02540*   COMPUTING THE INTERPOLATED TRACK POSITIONS.
02560 NROW=NPTS+NPTS
02580 DO 800 K=1,2
02600 NCOL=M(K)+M(K)
02620 DO 780 I=1,NROW
02640 TINT(K, I)=0.
02660 DO 760 J=1,NCOL
02680 TINT(K, I)=TINT(K, I)+(Q(K, I, J)+TR(K, J))
02700 760 CONTINUE
02720 780 CONTINUE
02740 800 CONTINUE
02760*
02780*   COMPUTATION OF SIG=Q+SG*(Q-TRANSPOSE)
02800 DO 900 K=1,2
02820 NCOL=M(K)+M(K)
02840 DO 880 I=1,NROW
02860 DO 860 J=1,NCOL
02880 SIG(K, I, J)=0.
02900 DO 840 N=1,NCOL
02920 DUM(I, N)=0.
02940 DO 820 L=1,NCOL
02960 DUM(I, N)=DUM(I, N)+(Q(K, I, L)+SG(K, L, N))
02980 820 CONTINUE
03000 SIG(K, I, J)=SIG(K, I, J)+(DUM(I, N)+Q(K, J, N))
03020 840 CONTINUE
03040 860 CONTINUE
03060 880 CONTINUE
03080 900 CONTINUE
03100 NCOL=NROW

```

HOWARD L. WIENER

```

03120*
03140*   COMPUTE TDIFF=TINT1-TINT2, DIFFERENCE VECTOR
03160 DO 920 I=1,NROW
03180 TDIFF(I)=TINT(1,I)-TINT(2,I)
03200 920 CONTINUE
03220*
03240*   COMPUTE SIGM=SIG1+SIG2, MATRIX SUM
03260 DO 960 I=1,NROW
03280 DO 940 J=1,NCOL
03300 SIGM(I,J)=SIG(1,I,J)+SIG(2,I,J)
03320 W2(I,J)=SIGM(I,J)
03340 940 CONTINUE
03360 960 CONTINUE
03380 PRINT 9050
03400 9050 FORMAT(/,20X,20H2.TRANSFORMED VALUES/19X,7HTRACK 1;
03420+17X,7HTRACK 2/3X,4HTIME,2X,2(5X,7HX-COORD,5X,7HY-COORD))
03440 DO 965 I=1,NPTS
03460 PRINT 9020,TCOM(I),TINT(1,2+I-1),TINT(1,2+I),TINT(2,2+I-1),
03480+TINT(2,2+I)
03500 965 CONTINUE
03520 PRINT 9060
03540 9060 FORMAT(/,15X,38HSUM OF TRANSFORMED COVARIANCE MATRICES)
03560 DO 970 I=1,NROW
03580 PRINT 9040,(SIGM(I,J), J=1,NPCOL)
03600 970 CONTINUE
03620*
03640*   COMPUTE R2=(TDIFF-TRANSPDSE)*(SIGM-INVERSE)*(TDIFF)
03660 CALL MATINV(W2,SIG1,16,32,W1)
03680 R2=0.
03700 DO 1000 J=1,NPCOL
03720 DUM(1,J)=0.
03740 DO 980 I=1,NROW
03760 DUM(1,J)=DUM(1,J)+(TDIFF(I)*SIG1(I,J))
03780 980 CONTINUE
03800 R2=R2+(DUM(1,J)*TDIFF(J))
03820 1000 CONTINUE
03840 PRINT 8000, R2
03860 8000 FORMAT(/,20X,17H3.COMPUTED VALUES,/3HR2=,F12.5)

```

NRL REPORT 7989

```

03880*
03900*  COMPUTE THE (1-ALPHA)-TH PERCENTILE POINT FOR THE
03920*  CHI-SQUARE DISTRIBUTION WITH 2*NPTS DEGREES OF FREEDOM
03940  XNPT=2*NPTS
03960  PROB=1.-ALPHA
03980  IF (PROB.LE.0.5) GO TO 1020
04000  P=1.-PROB
04020  GO TO 1040
04040  1020 P=PROB
04060  1040 TE=SQRT(ALOG(1./(P*P)))
04080  C0=2.515517
04100  C1=0.802853
04120  C2=0.010328
04140  D1=1.432788
04160  D2=0.189269
04180  D3=0.001308
04200  U1=1.+(D1*TE)+(D2*TE*TE)+(D3*TE*TE*TE)
04220  U2=C0+(C1*TE)+(C2*TE*TE)
04240  XP=TE-(U2/U1)
04260  XP=XP*(PROB-0.5)/(ABS(PROB-0.5))
04280  R=(XP*SQRT(2./(9.*XNPT)))+1.
04300  CHISQ=XNPT*((R-(2./(9.*XNPT)))+3)
04320  PRINT 8020,ALPHA,XNPT,CHISQ
04340  8020 FORMAT(6H1ALPHA=, F6.3, 2X, 15HDEG. OF FREEDOM=F4.0,
04360+2X,6HCHISQ=, F12.4)
04380*
04400*  DECISION POINT
04420  IF (R2-CHISQ) 1060,1060,1080
04440  1060 PRINT 8040
04460  8040 FORMAT(12HMERGE TRACKS)
04480  GO TO 2000
04500  1080 PRINT 8060
04520  8060 FORMAT(32HDO NOT MERGE TRACKS AT THIS TIME)
04540  GO TO 2000
04560  2000 CONTINUE
04580  STOP
04600  END

```

HOWARD L. WIENER

```
04620 SUBROUTINE ORDER(G,MM,H,N)
04640 DIMENSION Q(2,15),MM(2),H(15),XX(15)
04660 MMI=MM(1)
04680 DO 20 I=1,MM1
04700 XX(I)=Q(1,I)
04720 20 CONTINUE
04740 MMM=MM(1)+MM(2)
04760 MM2=MM(2)
04780 DO 40 I=1,MM2
04800 J=MM(1)+I
04820 XX(J)=Q(2,I)
04840 40 CONTINUE
04860 DO 80 J=1,MMM
04880 H(J)=XX(J)
04900 DO 60 I=J,MMM
04920 IF (XX(I).GE.H(J)) GO TO 60
04940 Z=H(J)
04960 H(J)=XX(I)
04980 XX(I)=Z
05000 60 CONTINUE
05020 80 CONTINUE
05040 N=MMM
05060 DO 120 I=1,MMM
05080 IF (H(I).NE.H(I+1)) GO TO 120
05100 I1=I+1
05120 DO 100 J=I1,MMM
05140 H(J)=H(J+1)
05160 100 CONTINUE
05180 N=N-1
05200 MMM=MMM-1
05220 120 CONTINUE
05240 MMM=MM(1)+MM(2)
05260 RETURN
05280 END
```

# memorandum

DATE: 5 January 1996  
Ser 96-018

REPLY TO  
ATTN OF: Code 8150

SUBJECT: REVIEW OF NRL REPORT 7989

TO: Code 1221

*AD - 3011 564*

1. It was requested that I review NRL Report 7989, "Track-to-Track Association in Ocean Surveillance", by Howard Weiner for a possible change in distribution statement. I have subsequently reviewed the copy provided and can see no reason for restricting access to this information.
2. It is recommended that the distribution statement on this subject report be changed to unlimited public release.

*R.L. Beard*

R.L. Beard  
Head, Space Applications Branch

*Completed*

*20 Mar 2000*

*B.W.*